DX
2997

UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK

FEDERAL HOUSING FINANCE AGENCY, AS CONSERVATOR FOR THE FEDERAL NATIONAL MORTGAGE ASSOCIATION AND THE FEDERAL HOME LOAN MORTGAGE CORPORATION, Plaintiff, -against- NOMURA HOLDING AMERICA INC., et al., Defendants.	No. 11-cv-6201 (DLC) ECF Case

AFFIDAVIT OF ARNOLD BARNETT

STATE OF NEW YORK)
) ss.:
COUNTY OF NEW YORK)

Arnold Barnett, being duly sworn, deposes and says:

1. I am the George Eastman Professor of Management and Professor of Statistics at the Sloan School of Management, Massachusetts Institute of Technology. I provide this affidavit as my direct testimony at trial. My testimony discusses proper approaches to achieving a random sample and drawing inferences from that sample, and explains why the sampling and extrapolation performed by plaintiff's sampling expert, Dr. Charles Cowan, is not reliable.

Professional Background

2. I earned a B.A. in Physics from Columbia University in 1969, and a Ph.D. in Applied Mathematics from the Massachusetts Institute of Technology ("MIT") in 1973.

3. I began my teaching career at MIT in 1971. In 1987 I was named a full Professor of Operations Research at the Sloan School of Management, and in 1998 I was named the George Eastman Professor of Management Science.

4. My research specialty is applied statistical analysis. During my career at MIT, I have taught seventeen different courses focusing on probability and statistics over a period spanning more than thirty-five years. These classes have covered everything from "Probability and Random Variables" to "Advanced Statistics" to "Advanced Calculus for Applications." I have also taught Biostatistics in the Harvard/MIT Clinical Investigators Training Program for nearly two decades.

5. I have received numerous teaching and research awards. For example, I am the only active academic ever to have received the President's Citation from the Flight Safety Foundation for "truly outstanding contributions on behalf of safety." I have been selected thirteen times for excellence in teaching recognition, most recently in 2013. In 2003 I was named a Fellow of the Institute of Management Sciences and Operations Research.

6. Between 1978 and 2013, I served as the principal investigator for twelve statistical research projects, sponsored by institutions ranging from the United States Department of Justice to the Federal Aviation Administration. These projects investigated a wide variety of topics, including airline safety and crime rates.

7. I have served as the editor for more than half a dozen publications, including *Management Science* and *Operations Research*.

8. A web course based on my Biostatistics class was created in 2006, and my textbooks on Applied Probability and Applied Statistics will both be published by March 2015. I have also published more than 100 articles and book chapters using statistical analysis to explore

a wide variety of topics and disciplines, with an emphasis on public policy and issues of health and safety.

9. Since 2010 I have provided expert witness testimony—either at trial or deposition—in 21 matters, including several that addressed residential mortgage-backed securities.

I. Principles Governing Sampling and Extrapolation

10. Sampling refers to the process by which one identifies and selects a subset of a larger population in order to learn about characteristics of the entire population. Extrapolation, in turn, refers to the process by which one generalizes information about the sample to the overall population. Together, sampling and extrapolation can permit one to make inferences about the properties of the entire population—without actually having to examine the entire population.

11. In the context of this litigation, the goal of sampling and extrapolation is to draw conclusions about the total population of loans underlying the at-issue securities based on analyzing a random sample of those loans. To perform this task, it was necessary (i) to use samples that are representative of the population, and (ii) to extrapolate properly from those samples.

12. One way to obtain a representative sample is by drawing a random sample. For a sample to be random, two criteria must be satisfied: (i) the selection of each sample unit is based on chance; and (ii) each unit of the population has a known, non-zero probability of being selected. After selecting a random sample, a statistician should confirm the representativeness of that sample. This is done by testing whether the characteristics of the members of the sample are similar to the characteristics of the population as a whole.

13. After selecting a random sample, a statistician can use information about the members of that sample to estimate characteristics of the overall population. He or she can also gauge the error rates of those estimates. Estimates based on extrapolation have error rates due to "sampling error," which refers to the fact that the sample is rarely a perfect replica of the full population. For example, if I were investigating the percentage of black marbles in a population of ten marbles (with five black and five white marbles) by choosing a random sample of two marbles, my sample could consist (in chronological order) of two white marbles, two black marbles, one white and then one black marble, or one black marble and then one white marble. Yet only two of these four possibilities would correctly reflect the 50/50 split that prevails in the full population. In the other two cases, the sampling error would be 50 percentage points, because each such case entails a 0/100 split.

14. Estimates based on extrapolation may also have error rates because of "modeling error," *i.e.*, error that is introduced when the characteristics of the sample are estimated using a model that itself is subject to error. For example, if I am estimating average height in a population based on a sample of individuals and using a device for measuring heights that is imperfect, then modeling error accompanies sampling error as a source of imprecision in the population-wide estimate.

15. An estimate for a statistic about an overall population typically consists of two components: (i) a point estimate, which represents the "best guess" for the population characteristic; and (ii) a confidence interval, which reflects the imprecision in the point estimate and provides a range of plausible values for the population characteristic. The confidence interval itself has two components: a confidence *level* (typically 95 percent), which specifies the probability that the confidence interval contains the true value of the population-wide

characteristic, and the margin of error, which is generally the distance from the point estimate to either the upper or lower bounds of the confidence interval. For example, if a sample-based point estimate of a particular characteristic is 25 percent, and the 95 percent confidence interval for that estimate extends from 20 to 30 percent, then the confidence level is 95 percent and the margin of error is 5 (30 – 25) percentage points. Correspondingly, the probability is 95 percent that the population characteristic is in the range of 20 to 30 percent. Outcomes within that range would be considered plausible given the sampling result, while outcomes outside the range would be considered less plausible (though not impossible, because by construction the population characteristic has a 5 (100 – 95) percent chance of falling outside the 95 percent confidence interval).

II. Case Background

16. In this case, plaintiff attempts to prove that the offering documents contained misrepresentations about the loan characteristics for loans underlying seven residential mortgage-backed securities by analyzing a sample drawn from the total population of loans comprising the supporting loan group for each security. To do this, plaintiff relies on the results reached by its sampling expert, Dr. Cowan.

17. In October 2012, Dr. Cowan proposed a sampling methodology for this case.[1] He proposed selecting a random sample of 100 loans from the total population of loans in each supporting loan group. According to Dr. Cowan, each sample was sufficiently large to enable him to make population-level estimates with a maximum margin of error of ±10 percentage points at a 95 percent confidence level.[2] Dr. Cowan also stated that he had tested the

---

[1] October 10, 2012 Expert Report of Charles D. Cowan.

[2] *Id.* at 2.

representativeness of his 100-loan samples and had determined that the samples were representative of the supporting loan groups respectively as a whole.[3]

18. Dr. Cowan's samples were then analyzed by two of plaintiff's other experts, Mr. Robert Hunter and Dr. John Kilpatrick. Mr. Hunter performed a re-underwriting analysis and purported to identify the number of "materially defective" loans in each sample, *i.e.*, loans that failed to comply with the originators' underwriting guidelines. Dr. Kilpatrick purported to determine the "true" value of each sample property using an automated valuation model he created, the Greenfield AVM ("GAVM"). Based on his GAVM results, Dr. Kilpatrick attempted with his Credibility Assessment Model ("CAM") to evaluate the credibility of those loans identified by the GAVM as having "inflated" appraisals.

19. However, neither Mr. Hunter nor Dr. Kilpatrick was able to evaluate all of the sample loans selected by Dr. Cowan.

20. Mr. Hunter stated that he was unable to evaluate 61 of the original 700 loans contained in the samples for the combined supporting loan groups. 53 of these missing loans were concentrated in the sample for the security called NAAC 2005-AR6. To replace the loans Mr. Hunter claimed could not be re-underwritten, Dr. Cowan drew 96 supplemental sample loans from the supporting loan group for that security. Mr. Hunter ultimately re-underwrote 131 loans from the combined initial and supplemental samples drawn for NAAC 2005-AR6. He stated that he could not re-underwrite 65 of the 196 total loans for that group, a shortfall of 33 percent of the combined samples.[4] Dr. Cowan, however, never tested the representativeness of the final sample of 131 loans from NAAC 2005-AR6 that Mr. Hunter re-underwrote. Mr. Hunter

---

[3] *Id.* at 11.

[4] May 15, 2014 Report of Robert W. Hunter ("Hunter Report") at 2 n.4.

also stated that he was unable to re-underwrite 8 of the loans contained in the samples for the remaining securities.[5]

21. Similarly, Dr. Kilpatrick stated that he was unable to provide reliable GAVM values for 15.6 percent of the loans in the samples drawn by Dr. Cowan.[6] DX 2824 shows the number of loans from each sample that Dr. Kilpatrick could not value using the GAVM. Again, Dr. Cowan never tested whether the final samples that Dr. Kilpatrick valued using the GAVM were representative of the supporting loan groups from which the samples were drawn.

22. Dr. Cowan purports to extrapolate these sample-level results. With respect to Mr. Hunter's analysis, Dr. Cowan estimates both count-based materially defective rates (fractions of loans with defective underwriting) and dollar-weighted materially defective rates (ratios of the dollar value of defective loans to the dollar value of all loans, as measured by the loans' original principal balances). He also extrapolates Mr. Hunter's various related findings (for example, the percentage of loans for which Mr. Hunter claims occupancy was misstated).[7]

23. Dr. Cowan purports to extrapolate Dr. Kilpatrick's analysis, as well. In so doing, he states that he estimates the degree of appraisal inflation for loans in each supporting loan group; the percentage of loans with allegedly understated LTV ratios at the 95 percent confidence level; and the percentage of loans with non-credible appraisals.[8]

---

5 *Id.*

6 October 7, 2014 Report of Charles D. Cowan ("Cowan Report") at 9-10.

7 *Id.* at 4-9.

8 *Id.* at 9-19.

24. Dr. Cowan's sampling and extrapolation methodologies contain serious flaws. The final samples used by Mr. Hunter and Dr. Kilpatrick are non-random, and Dr. Cowan's extrapolations of Mr. Hunter's and Dr. Kilpatrick's results are unreliable and prone to systematic bias. As a result, his ultimate conclusions are highly unreliable.

25. Defendants' expert Mr. Michael Forester also analyzed the sample loans selected by Dr. Cowan. He re-underwrote those loans identified as materially defective by Mr. Hunter and concluded that Mr. Hunter drastically overstated the material defect rate. To maximize comparability with Dr. Cowan's results, I extrapolate Mr. Forester's results by generally applying Dr. Cowan's methods.

III. If Dr. Kilpatrick's and Mr. Hunter's Results Are Unreliable, Dr. Cowan's Extrapolations of those Results Are Likewise Unreliable

26. Dr. Cowan purports to do nothing beyond extrapolating Dr. Kilpatrick's and Mr. Hunter's conclusions from the sample to the general population. Therefore, by definition, Dr. Cowan's results can only reliably estimate statistics about the supporting loan groups if the corresponding statistics within the samples were accurately measured.

27. If plaintiff's other experts failed to make accurate assessments about the samples, then Dr. Cowan's extrapolation methods will only serve to propagate those errors. For example, if Dr. Kilpatrick's GAVM is unreliable, then Dr. Cowan's extrapolation of those results to an entire supporting loan group is similarly unreliable.

IV. For Securitizations In Which Mr. Hunter's Samples Were Not Shown to Be Random and Representative, Extrapolations of Mr. Hunter's Results Are Unreliable

28. Dr. Cowan did not show that the samples from which he extrapolates Mr. Hunter's final results were random and representative: he unaccountably failed to conduct the representativeness tests that he applied to his initial samples for the subsequent, final samples

that were actually analyzed by Mr. Hunter. Even when there are appreciable differences between the initial and final samples, Dr. Cowan offers no evidence that the final extrapolations for these samples are reliable.

29. To perform an accurate extrapolation using Dr. Cowan's approach, the final sample must be randomly selected. If random selection does not occur (for example, if loans with certain characteristics have no chance of being included in the sample), then the results from the sample cannot reliably be extrapolated to the full population. While Dr. Cowan's initial samples appear to have been selected randomly, Dr. Cowan does not show that the final samples—the only samples that matter for purposes of the ultimate extrapolation—were not created under a non-random process of deletion.

30. Mr. Hunter claimed not to be able to re-underwrite many loans in the initial samples because they lacked adequate documentation. Overall, Mr. Hunter failed to re-underwrite 61 loans (8.7 percent) of Dr. Cowan's initial samples of 700 loans. Fifty-three of these loans were concentrated in the supporting loan group for NAAC 2005-AR6. All groups contained at least one missing loan, with the exception of NHELI 2006-FM1 and NHELI 2006-FM2.

31. Dr. Cowan tested none of the final samples for representativeness. This omission is particularly acute for NAAC 2005-AR6, for which Mr. Hunter excluded 65 loans chosen for the original and supplemental samples. When there is no evidence that the final samples are representative, there is no basis to conclude that Dr. Cowan's extrapolations based on those final samples are reliable

32. Compounding this problem is the fact that Mr. Hunter's criteria for rejecting a loan due to an inadequate loan file have not been disclosed. In some cases, he

apparently considered a loan file to be incomplete (and thus incapable of being re-underwritten) if it was missing certain documents, while in other cases he proceeded to re-underwrite a loan file missing the same documents and found it defective.[9]

33. If the loans that Mr. Hunter was unable to re-underwrite shared a common characteristic, that characteristic would, as a result, have been underrepresented in the final sample. For example, suppose that loans secured by properties in New York were unusually common among the missing loans. Because the supplemental sample was chosen in the same fashion as the initial one, it would likely reflect a similar deficit. In consequence, replacing absent loans secured by New York properties with new loans randomly selected might effectively substitute loans secured by properties in California or Florida. The upshot is that the final sample would suffer from a geographic bias even though both the initial sample and the supplemental sample were randomly selected. Similarly, if there were few complete loan files available from a particular originator, the loans that replaced those initially sampled from that originator would tend to come from other originators, thus perpetuating the bias.

34. A selection bias will undermine any extrapolation process dependent on random sampling if the loans Mr. Hunter actually re-underwrote differed in salient respects from loans comprising the broader population. For example, if loans that were prepaid due to refinancing were less likely to have complete loan files because they had been closed for a longer period of time than other loans, they likely would be underrepresented in the final sample. If prepaid loans were also less likely to be found defective, the percentage of defective loans found by Mr. Hunter would be overstated, and Dr. Cowan's extrapolated materially defective rate for that supporting loan group would likewise be overstated.

[9] Hunter Report at 90-91.

35. Lacking evidence that a final sample was representative in all salient respects of a supporting loan group's general population, we likewise lack evidence that Dr. Cowan's extrapolation of Mr. Hunter's results for that supporting loan group is reliable.

V. Because Most of Dr. Kilpatrick's Samples Were Not Shown to Be Random and Representative, Extrapolations Based on Dr. Kilpatrick's Results Are Unreliable

36. Dr. Kilpatrick provided GAVM values for only subsets of the loans in Dr. Cowan's initial samples. In total, Kilpatrick failed to provide GAVM estimates for 110 (15.7 percent) of the 700 loans included in the original samples that Cowan drew. He did not evaluate 100 percent of the sample loans for any loan group.

37. In four of the loan samples for which Dr. Kilpatrick obtained GAVM results, the truncation of the sample is a particularly serious concern. Dr. Kilpatrick provided GAVM values for fewer than 90 percent of the sample loans in four of the seven supporting loan groups, *i.e.*, those for NAAC 2005-AR6, NHELI 2006-HE3, NHELI 2007-2, and NHELI 2007-3. For those four supporting loan groups, Dr. Kilpatrick failed to provide GAVM estimates for 91 (22.75 percent) of the 400 loans included in the initial samples. Of the 496 loans included in the combined initial and supplemental samples for these four securities, Dr. Kilpatrick failed to provide estimates for 105 loans, or 21.17 percent.

38. Again, Dr. Cowan never evaluated the representativeness of the final samples evaluated by Dr. Kilpatrick, and thus failed to show that his extrapolations of Dr. Kilpatrick's results for those samples are reliable.

VI. Dr. Cowan Failed to Extrapolate "Defect" Rates By Originator

39. Apart from the problem of their non-representativeness, Dr. Cowan's samples are also too small reliably to extrapolate defect rates at the originator level for the five securitizations in which multiple originators were disclosed in the prospectus supplements. Dr.

Cowan has not attempted to make such an extrapolation. Extrapolated defect rates on an originator-by-originator basis would yield margins of error so large that they would not only exceed Dr. Cowan's ±10 percentage point threshold discussed in paragraph 17, above, but would also render defect-rate estimates for individual originators so imprecise as to be meaningless.

VII. Additional Problems with Dr. Cowan's Extrapolation of Dr. Kilpatrick's GAVM Results

40. Dr. Cowan's extrapolations of Dr. Kilpatrick's GAVM results are highly problematic for additional reasons, too. He makes several methodological choices that result in a systematic upward bias in his LTV recalculations.

41. To extrapolate Dr. Kilpatrick's conclusions, Dr. Cowan performed a Monte Carlo simulation, which is designed to take into account two forms of variability: the variability that accompanies the use of a sample and the variability inherent in Dr. Kilpatrick's GAVM, which produces not just a point estimate for the "true" property value but also a probability distribution for that quantity. This distribution is intended to account for GAVM model error. Dr. Cowan assumes that the probability distribution for the "true" property value is normal, or bell-shaped, with its mean equal to the GAVM point estimate and a standard deviation (labeled a "forecast standard deviation") provided by the GAVM that determines the width of the bell-shaped curve.

42. Dr. Cowan uses the simulated values produced by the Monte Carlo simulation and based on the GAVM to calculate average appraisal inflation rates and to recalculate LTV ratios for the loans underlying the at-issue securities.

a. Dr. Cowan Ignores Dr. Kilpatrick's Definition of an Inflated Appraisal

43. In estimating average appraisal inflation per supporting loan group, Dr. Cowan deviates from Dr. Kilpatrick's definition of an "inflated" appraisal.

44. Dr. Kilpatrick defines an appraisal as inflated (or undervalued) only if it is more than one standard deviation above (or below) the corresponding GAVM value. Dr. Cowan claims to accept Dr. Kilpatrick's definition. Yet in performing his average inflation rate calculations, Dr. Cowan relies exclusively on the simulated values produced by his Monte Carlo simulation, even when the original appraisal was within one standard deviation of the GAVM value. Had he been faithful to Dr. Kilpatrick's convention, Dr. Cowan would have used an inflation rate of zero for all loans that did not satisfy Dr. Kilpatrick's standard for inflation.

45. DX 2825 reflects my recalculations of average inflation per supporting loan group using Dr. Kilpatrick's standard. These revisions yield lower estimates of appraisal inflation rates for every single securitization. The recalculated aggregate inflation rate is under 10 percent (9.9 percent).

46. The revision affects Dr. Cowan's LTV recalculations as well. In determining the "value" segment of the loan-to-value ratio, Dr. Cowan would have been consistent with Dr. Kilpatrick's standard if he used the original appraisal values for all loans that did not qualify as inflated under that standard.

b. Dr. Cowan's LTV Calculations Suffer from a Serious Upward Bias

47. In recalculating LTV ratios for the loans for which Dr. Kilpatrick provided GAVM estimates, Dr. Cowan follows a procedure that generates a systematic upward bias in his results. This is true for three distinct reasons.

48. First, even if one treats the GAVM point estimates as correct on average, imprecision in the GAVM results leads to an upward bias when using those valuations in the denominators of LTV calculations. The GAVM point estimate for any particular property is just as likely to be too low by a particular amount as it is to be too high by that same amount. In the context of LTV calculations, these opposite variations are not offsetting. Instead, a point

estimate that is too low raises the LTV ratio by more than an estimate that is too high by the same amount reduces the ratio. The upshot is that even a GAVM that generates unbiased point estimates in the aggregate nonetheless would produce artificially high estimates of the LTV ratios.

49. An example illustrates this effect. Suppose that one is dividing 12 by X, and assume further that the "true" value of X is 6. But suppose that an estimate of X is not exactly 6 but is instead equally likely to take on three different values that average out to 6: 4, 6 and 8. The average of the 12/X ratios assuming those three possible values for X—((12/4 = 3); (12/6 =2); (12/8 =1.5))—is 2.17 (*i.e.*, (3 + 2 + 1.5)/3 = 2.17). That result is 8.5 percent higher than the true ratio (12/6 =2). In the same way, the random error in GAVM point estimates has the effect of producing LTV estimates that skew higher than the correct LTV even if the GAVM values standing alone are unbiased. DX 2826 illustrates this effect.

50. Dr. Cowan did not actually use the GAVM point estimates as the source of his LTV recalculations. Instead, he adopts a procedure even more unreliable, which increases the upward bias in his recalculated LTV ratios. His Monte Carlo simulation compounds the bias in the GAVM point estimates by selecting random values within the GAVM's probability distribution to reflect uncertainty about what the "true" property value is. Half of the time, the simulation selects an estimated valuation that is lower than the GAVM's point estimate, and the other half of the time, it selects a valuation higher than the point estimate. But by the same reasoning as in the previous paragraph, low-end valuations push the LTV ratio estimate up by more than the equally likely high-end valuations push the estimate down. Thus, the average LTV calculated on the basis of the Monte Carlo simulation exhibits more upward bias than would use of the point estimate itself.

51. Dr. Cowan increases the upward bias by yet a third aspect of his valuations for individual properties. He recalculates the LTV ratio for each loan using the original principal balance divided by the minimum of the original appraisal value, the actual sale price (for purchase loans only), or the simulated GAVM. The effect of this asymmetric procedure is to treat valuations arising from the GAVM as reliable only when they fall below appraisal values for refinance loans or fall below appraisal values and sales prices for purchase loans; otherwise, the GAVM valuations are discarded. Such a procedure is indefensible under Dr. Cowan's assumption that the GAVM point estimates are on average correct. Indeed, Dr. Cowan's LTV analysis is self-contradictory. It is literally impossible under Dr. Cowan's method to conclude that even a single loan-tape LTV ratio was overstated, even if the GAVM valuation strongly implies that it was.

52. DX 2827 illustrates the consequences of using Dr. Cowan's biased procedure. It contrasts the loan-tape LTV ratios, LTV ratios calculated on the basis of Dr. Kilpatrick's GAVM point estimates, and Dr. Cowan's recalculated LTV ratios. As DX 2827 shows, a large percentage of the LTV ratios calculated using Dr. Kilpatrick's GAVM point estimates (the green dots) fall below the loan-tape LTV ratios (the red line), indicating that these ratios based on point estimates are lower than the loan-tape LTV ratios. In contrast, Dr. Cowan's recalculated LTV ratios fall—as they must—only on or above the red line. This pattern reflects a systematic upward skew in Dr. Cowan's results that cannot be reconciled with the GAVM that he purports to treat as unbiased.

c. Dr. Cowan's Results Regarding LTV Ranges Depend Heavily on Differences that He Himself Describes as Undetectible

53. Finally, Dr. Cowan uses a dubious method in recalculating the percentage of loans that fall into each LTV range (for example, the percentage of loans with an LTV ratio between 80 and 85, or over 100).

54. Dr. Cowan's first step in producing his LTV ranges involves calculating a 95 percent confidence interval for each loan's recalculated LTV ratio, *i.e.*, the range of plausible values for the "true" LTV produced by his own (already biased) method. He finds that the loan-tape LTV ratios fall below the confidence interval for 40 of the 672 loans (about six percent) in the sample.[10] For the remaining 632 loans, Dr. Cowan reports that there is "no detectible difference" between the recalculated LTV ratio and the loan-tape LTV ratio. [11]

55. This six percent estimate for understated LTV ratios is itself too high. Even if Dr. Cowan's method was completely unbiased and all loan-tape LTV ratios were accurate, 2.5 percent of these ratios would fall below the 95 percent confidence interval lower bound. This happens because, when the confidence interval has a 95 percent chance of including the "true" LTV ratio, there is a 5 percent chance of it not doing so. That 5 percent is divided between a 2.5 percent chance that the "true" LTV ratio falls above the confidence interval and a 2.5 percent chance that it falls below. Thus, Dr. Cowan's 6 percent rate of understated LTV ratios should be contrasted with a 2.5 percent baseline rate that would have been expected even for "true" LTV ratios. The actual rate of understated loan-tape LTV ratios even within Dr.

---

[10] Cowan Report at 14.

[11] Cowan Tr. at 252:20-253:4.

Cowan's own framework would thus have been far closer to 3.5 (6 – 2.5) percent than to 6 percent.

56. For purposes of determining the percentage of loans that fall into each LTV range, Dr. Cowan ignores his own conclusions about detectible differences and accepts literally hundreds of differences that were not "detectible." Had he instead replaced the loan-tape LTV ratios only in the 40 instances that fell below his confidence intervals, his conclusions about LTV ranges would have been dramatically different. DX 2828 makes this outcome clear. The column "Cowan at 95% Confidence Level" displays the LTV ranges that result when one uses loan-tape values for all loans other than the 40 that Dr. Cowan concluded had LTV ratios that were understated to a statistically significant degree.

VIII. Dr. Cowan's Extrapolation of Dr. Kilpatrick's Credibility Results Applies Only to Appraisals Found to be "Inflated" by the GAVM—Not to All Appraisals

57. Dr. Kilpatrick assessed the so-called "credibility" of loan-tape appraisals only for a small subset of the sample loans. As a result, his credibility conclusions are not generalizable to the total population.

58. Dr. Kilpatrick relies on his CAM to make what he contends is a determination about whether reasonable appraisers could have believed the opinions that they delivered about value at the time loans were originated. He performed this analysis only with respect to appraisals that he defined as "inflated" based on his GAVM results. He concluded that approximately 92 percent of the appraisals that he considered inflated were not credible.[12]

59. Because Dr. Kilpatrick performed a credibility analysis on only that subset of the sample loans he found to be inflated under his "one standard deviation" rule, he ultimately

[12] Cowan Report at 16.

provided no credibility assessments for 73.3 percent of the sample loans. It is inappropriate to draw credibility conclusions about all loans in the general population based on such extremely limited results, as Dr. Cowan himself agrees.

IX. Considered Together, the Results of Dr. Kilpatrick's GAVM and CAM Call Into Question the Reliability of Both, and Dr. Cowan Fails to Account for This Problem

60. Dr. Kilpatrick's CAM and GAVM results are in serious tension with each other. As I explain below, many loans that Dr. Kilpatrick classifies as non-credible are, in fact, not actually inflated under the probability distributions arising from his GAVM. As a result, Dr. Cowan's extrapolations of these contradictory results is highly problematic.

61. Dr. Cowan concludes, based on Dr. Kilpatrick's sample-level data, that approximately 92 percent of the appraisals for "inflated" loans at the population level were non-credible.

62. However, as Dr. Cowan admitted at his deposition, even if all the original appraisals (conducted at the time of the origination of the loans) were correct *and* the GAVM is reliable, 16 percent of the original appraisals would still meet Dr. Kilpatrick's criterion for inflation due to random chance alone.[13] This effect occurs because, under the GAVM's normal probability distribution for the "true" property value, there is a 16 percent chance that the correct valuation will be more than one standard deviation above the GAVM point estimate. Thus, the GAVM likely would have produced approximately 107 inflated appraisal "false positives" (16 percent of the 672 loans for which Kilpatrick generated a GAVM value), even if all appraisals were exactly correct.

[13] November 20, 2014 Deposition of Charles D. Cowan at 274-75.

63. Dr. Kilpatrick designated as inflated 199 of the 672 loan-tape appraisals that he considered. But if 107 of the 199 were false positives, then the number of true positives would be 92 (199 – 107). For technical reasons, a better estimate of the number of false positives may be 90 rather than 107.[14] But that figure still implies that roughly half of the 199 "inflated" valuations were correct under Dr. Kilpatrick's own model.

64. We have, then, an apparent contradiction. Approximately half the "inflated" appraisals were accurate under the GAVM, yet only 8 percent of these appraisals were deemed credible (100 – 92 = 8).

65. This sharp inconsistency indicates that either the GAVM or the CAM is unreliable, or that both of them are. That Dr. Kilpatrick's GAVM values and his credibility determinations are incompatible casts grave doubt both on his results and Dr. Cowan's extrapolation of those results.

X. Dr. Cowan's "Defect" Rate is Questionable Because of a Distinction among Originators That Was Not Considered in the Re-Underwriting Process

66. I understand that defendants have argued that the prospectus supplements make different representations about originators whose identity and underwriting guidelines were disclosed in the prospectus supplements ("disclosed originators"), and those originators not so disclosed (the "undisclosed originators"). Defendants contend that the criteria under which the re-underwriters evaluated loans from undisclosed originators were inappropriate.

67. If Mr. Hunter used the wrong criteria for determining whether a given loan was defective, then extrapolations based on his re-underwriting results are necessarily unreliable. It could be argued that, if plaintiff had evaluated loans from undisclosed originators in an

[14] This number is based on the approximate equation $x + 0.16(672 – x) = 199$, where x is the number of true positives.

improper way, then the defects they allegedly find among such loans should not be counted in the defect-rate calculation. Under that reasoning, Mr. Hunter's defect rate estimates would be too high, as would Dr. Cowan's extrapolations based on these estimates.

XI. Dr. Cowan Offers Incorrect Margins of Error for His Dollar-Weighted Defect Rates

68. Dr. Cowan's extrapolation of dollar-weighted materially defective rates based on Mr. Hunter's results is flawed, because it relies on an improper extrapolation methodology.

69. To calculate the margin of error for his extrapolation of dollar-weighted materially defective rates, Dr. Cowan uses a method based on binominal distribution, which is inappropriate in this context. The binomial method weights all loans equally, but dollar-weighted estimates, by definition, weight larger loans greater than smaller ones. Thus, Dr. Cowan's margins of error, 95 percent lower bounds, and 95 percent upper bounds of his dollar-weighted materially defective rates are incorrect.

70. I recalculated the margins of error for the dollar-weighted materially defective rates using a method based on bootstrapping. My results are reflected in DX 2829. I find that the margins of error for the dollar-weighted material defect rates calculated using this approach exceed Dr. Cowan's reported margins of error for four of the seven supporting loan groups.

XII. Extrapolation of Mr. Forester's Results

71. I also extrapolated the results of Mr. Forester's re-underwriting analysis from the samples to the associated populations. This extrapolation results in population-level potential defect rates that are far lower than the material defect rates calculated by Mr. Hunter and Dr. Cowan. I generally followed Dr. Cowan's conventions for extrapolation, because

introducing any differences between our extrapolation methodologies would confuse the comparison between the findings of Mr. Forester and Mr. Hunter.

72. Mr. Forester re-underwrote the loans that Mr. Hunter determined were materially defective. He ultimately concluded that 40 of the 723 sampled loans across the seven supporting loan groups might potentially have substantial defects.

73. DX 2830, attached, shows the count-based results of my extrapolation of Mr. Forester's "potential substantial defect" rate. As noted, I adopted Dr. Cowan's stratification scheme and formulas to maximize comparability with Dr. Cowan's extrapolation results. The extrapolated potential substantial defect rates range from 3 percent in one supporting loan group to 10 percent in another. On average across the seven supporting loan groups in this action, the extrapolated potential defect rate is 5.46 percent, with an overall margin of error of fewer than two percentage points. This rate is significantly lower than Mr. Hunter's extrapolated defect rate of 68.56 percent.

Arnold Barnett

Arnold Barnett

Sworn before me this
20 day of February 2015.

Notary Public

LILLIAN I. AYALA
Notary Public
COMMONWEALTH OF MASSACHUSETTS
My Commission Expires
February 23, 2018